



TAB

15 JUL 1972

4.8 DEFINE DICTIONARY REQUIREMENTS

Introduction

A dictionary is a file (or "data list", in GIM terminology) that consists of simple code-text relationships. The code makes it possible to prevent storing larger attributes many times in a data file, but to store them only once as "data" to a dictionary file. An example would be to set up a state code dictionary so that all employees from Alabama would not have to have "Alabama" coded as part of their personnel file but would need only to have a code such as "Ø1" stored instead. For report purposes, a program would only need to reference the state code dictionary to translate the "Ø1" to "Alabama".

Criteria for block changes may be established at a future time, once the software capabilities of GIM are known. At this time, block changes will not occur on a fully automatic basis. Co-ordination in the Data Management Center will be required with the projects involved to insure that dictionaries and user files are updated concurrently. The controlling project will coordinate an update for each attribute documented by the submitting project.

It is conceivable that a dictionary could have more than one text for the same code. Such a situation exists on the M&P Dictionaries today. However, the simple case will be assumed throughout this section, unless an indication is made otherwise.

Terms:

Controlling Project (CP): The project having responsibility for co-ordinating the set-up and maintenance of dictionaries. In HRS, this is Project HUMCO.

Submitting Project (SP): The project that submits the documentation of needed dictionaries to the CP. The requirement to submit documentation is not just for the project that will be requesting updates. It includes any user that needs to be considered in the design and maintenance of the dictionary. The

documentation includes reporting and maintenance requirements, dictionary characteristics, and data.

There may be more than one project that is a SP for a dictionary; any project having a vital interest in the establishment of a dictionary should submit its written requirements to the CP. If the requirements of two SP's are incompatible and can't be resolved, two separate dictionaries may be established. If the requirements of two SP's are compatible, one dictionary may be established to serve them both.

User Projects (UP): Those projects which will use the dictionaries but will not necessarily control the data in them. A UP that submits no documentation may not get a dictionary to use if no other project has submitted documentation.

The outline (Table of Contents) to be used in defining dictionary requirements is:

- A. Project Identification
- B. Active File Requirements (Form 2968)
- C. Maintenance Requirements
- D. Reporting Requirements
- E. Data Listing
- F. Data Editing
- G. Historical File Requirements

The following paragraphs explain the outline and the type of information initially required from the user/analyst. The controlling project (Project HUMCO) will use this to develop all dictionaries needed in the Human Resources Systems.

A. PROJECT IDENTIFICATION:

Identify your project by name.

OPTIONAL: If the dictionary already exists in one form or another (a dictionary that already has the codes needed is an excellent candidate even if it lacks the text), identify it. Whether or not the dictionary has already been established, you are aware of other projects which may use it in the future, list them too.

An analyst might identify a state-code dictionary his project needs in the following manner.

This dictionary is to be used by PERSIGN. The code and texts now exist as parts of the dictionary HRSSTATE.

B. ACTIVE FILE REQUIREMENTS (Form 2968):

The following information should be supplied for each code and text the SP intends to use: Field Name, Minimum (optional) and Maximum Field Lengths, Type (A, B, N, or X as defined in sub-step 4.8F), Justification (L or R), and Field Description. If a position by position character description is desired, it may also be described using the substep 4.8F as a guide.

The Field Description should include what is to be done with the data (search, retrieve...), why it is to be done (for reports, for computation, for validation...), and any limitations to be imposed on input or output. A sample Field Description might specify "This text is to be retrieved for logical control purposes. It is a code set up by PERSIGN to edit state income tax data items of the related state. The value is to be available to PERSIGN and PAYROLL only".

The Field Name supplied will be used when the dictionary is established. The name will be modified only to the extent that the Field Name's prefix will be replaced by the prefix belonging to the controlling project. For example, suppose project CAPER/OP has an attribute COPSTATECOD in its file that needs to be cross-referenced to a code and text in a dictionary it documents as COPSTATECOD. If the code and text are documented as COPSTATECOD and COPSTATETXT, the CP (HUMCO) will rename the fields to HRSSTATECOD and HRSSTATETXT respectively.

In the UP's (user project) 2968 for COPSTATECOD a notation should be made that the text is stored on the dictionary HRSSTATECOD. Through the proper GIM spans set up by the user, when COPSTATECOD is referenced the text will be retrieved automatically and identified as HRSSTATETXT.

C. MAINTENANCE REQUIREMENTS:

This section should tell the DMC who is authorized to make the requests and how the requests will be made. It will be assumed that Form 930 will be used to initiate the action and signed by the requesting component.

If the request will be via other than a 930, that method should be spelled out.* The component should be identified at the lowest organizational level known for sure at the time the documentation is prepared (in the early stages, this might be at the Office level).

* For example - internal updating of a data list would also update a dictionary file. e.g., ORGANIZATIONAL CODE/TEXT.

D. REPORTING REQUIREMENTS:

This section should specify the information to be published for the user projects when the dictionary is updated. It is anticipated that BARF (Basic Automatic Report Formatter) will produce the reports. The following information is a minimum requirement: if BARF is not used a spec sheet will require more.

What data (attributes) the report should contain, how many copies should be made, how often it should be published, and to whom the report should be sent.

- Specification examples:
- (a) The codes in the dictionary will be republished with each update. Send one copy to the Identification Section in O/S.
 - (b) The entire dictionary is to be republished quarterly. Send one copy each to O/F and O/S; two copies to O/P.

E. DATA LISTING:

If the dictionary does not yet exist, give the complete selection of codes and texts needed. That is, for each Form 2968 submitted, a corresponding code or text should be included. The listing may be merely a listing of an OS/360 computer file (Recommended: a dictionary pre-established on the M&P dictionary system).

At this time, there are two known character restrictions on data. No way has been found to have GIM accept a double quote mark or an asterisk (" or *): these are key characters in the GIM language.

F. DATA EDITING (optional):

On the 2968's filled in substep 4.8B, a specific pattern may be used for character by character editing. These may consist of the first six key characters illustrated, a character redefined with "\$" as illustrated, or any legal character as defined on the next page.

- * E = Empty: no prescribed value pattern from that direction
- ; = Right and left pattern separator
- * A = Alphabetic character only
- * B = Alphabetic including blank
- * N = Numeric character only
- * X = Any character (blank included)
- \$ = Redefines E, ;, A, N, X, B, :, ', /, or \$ as a literal when immediately preceding it to make it a legal character.

in the relative
character position
of the value.



A legal character is any literal other than E, ;, A, N, X, B, :, ', /, ", *, blank

Sample patterns:

'ABNNN;E'

From left to right:

- First character must be alphabetic.
- Second character must be alphabetic.
- Next three characters must be numeric.
- No specifications from right to left.

'\$A\$BXXX;E'

From left to right:

- First character must be letter A.
- Second character must be letter B.
- Next three characters may be any legal char.
- No specifications from right to left.

'\$A\$B;E'

Same result as above example.

'E;NNN.NN'

No specifications from left to right.

From right to left:

- First two characters must be numeric.
- Following character must be a decimal.
- Next three characters must be numeric.

G. HISTORICAL FILE REQUIREMENTS:

A "no" will suffice if the SP anticipates no need for same. If the answer is "yes" these questions should be answered:

1. Is the use of the historical data to be any different from the field description "active" data in substep 4.8B?
2. Once a historical item has been inactive for a given length of time, is it to be deleted? If yes, how long should that time be?
3. If the useage is for retrieval, will be requirement cover just the text for the date of the record?

all the texts that have ever been valid?

all the texts valid during a given range of dates?

A sample historical requirement description might be:

- to provide a record of all the texts (PERORGTEXT) that have ever been valid for a particular code (PERORGCODE),
- to provide only the text in effect at a given date, and
- automatic conversion of a historical code using the date of the historical record to provide the proper text.